DOCUMENT RESUME

ED 071 618 HE 003 757

AUTHOR	Hind, Robert E.; And Others
TITLE	The Evaluation of University Teachers: An Application of a Theory of Evaluation and Authority.
INSTITUTION	Stanford Univ., Calif. Stanford Center for Research and Development in Teaching.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	SCRDT-TR-32
BUREAU NO	BR-5-0252
PUB DATE	Dec 72
CONTRACT	OEC-6-10-078
NOTE	22p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Evaluation Methods; *Faculty Evaluation; *Higher Education; *Performance Criteria; Research; *Teacher Evaluation; Teaching

ABSTRACT

This paper reports on an investigation of the evaluation of teaching and research at Stanford University in the School of Humanities and Sciences. It was found that professional colleagues had far more influence on organizational rewards and penalties than did administrative superiors. Evaluations of research had far more influence than did evaluations of teaching. Since the faculty expended considerable effort on teaching, many wanted a better balance between effort and reward. For teaching, students were influential evaluators, since teaching performances were more visible to them. For research, colleagues in other institutions were influential evaluators, since the visibility of published research made the boundaries of the University permeable. Overall, for both teaching and research, department colleagues and department heads were the most influential evaluators. Faculty members' satisfaction with the evaluation process was positively associated with the perceived dependence of their discipline on theory. Greater dependence on theory in a discipline was associated with more perceived agreement among evaluators, and more agreement among evaluators was positively related to satisfaction with the evaluation system. (Author/HS)

FILMED FROM BEST AVAILABLE COPY

ED 071613

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

HE 003 557

SCRDT

Stanford Center for Research and Development in Teaching

SCHOOL OF EDUCATION STANFORD UNIVERSITY

STANFORD CENTER
FOR RESEARCH AND DEVELOPMENT
IN TEACHING

ED 071618

Technical Report No. 32

THE EVALUATION OF UNIVERSITY TEACHERS: AN APPLICATION OF A THEORY OF EVALUATION AND AUTHORITY

Robert R. Hind, Sanford M. Dornbusch,
and W. Richard Scott

School of Education
Stanford University
Stanford, California

December 1972

Published by the Stanford Center for Research and Development in Teaching, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred. (Contract No. OEC-6-10-078, Component 2B.)

Introductory Statement

The Center's mission is to improve teaching in American schools. Too many teachers still employ a didactic style aimed at filling passive students with facts. The teacher's environment often prevents him from changing his style, and may indeed drive him out of the profession. And the children of the poor typically suffer from the worst teaching.

The Center uses the resources of the behavioral sciences in pursuing its objectives. Drawing primarily upon psychology and sociology, but also upon other behavioral science disciplines, the Center has formulated programs of research, development, demonstration, and dissemination in three areas. Program 1, Teaching Effectiveness, is now developing a Model Teacher Training System that can be used to train both beginning and experienced teachers in effective teaching skills. Program 2, The Environment for Teaching, is developing models of school organization and ways of evaluating teachers that will encourage teachers to become more professional and more committed. Program 3, Teaching Students from Low-Income Areas, is development materials and procedures for motivating both students and teachers in low-income schools.

The research reported here was conducted in association with the studies of evaluation and authority within the Environment for Teaching program.

Abstract

This paper reports an investigation of the evaluation of teaching and research at Stanford University in the School of Humanities and Sciences. All the data were collected in 1968 from interviews of a sample of 100 faculty, stratified by rank.

It was found that in this professional bureaucracy, professional colleagues had far more influence on organizational rewards and penalties than did administrative superiors. Evaluations of research--by all evaluators combined--had far more influence than did evaluations of teaching. Since the faculty expended considerable effort on teaching, many wanted a better balance between effort and reward. For teaching, students were influential evaluators, since teaching performances were most visible to them. Information volunteered by the faculty indicated that they were divided about the quality of student evaluations. For research, colleagues in other institutions were influential evaluators, since the visibility of published research made the boundaries of the University permeable. Younger faculty, more mobile, emphasized research more than older faculty did. Overall, for both teaching and research, department colleagues, including department heads, were the most influential evaluators.

Faculty members' satisfaction with the evaluation process was positively associated with the perceived dependence of their discipline on theory. Greater dependence on theory in a discipline was associated with more perceived agreement among evaluators, and more agreement among evaluators was positively related to satisfaction with the evaluation system. This finding appears to be the basis for the professional emphasis on collegial evaluations, that is, evaluations based on a shared body of knowledge.

The study reported here provides the baseline for a current study of the effect of attempts made since 1968 to increase the weight given to teaching at Stanford.

THE EVALUATION OF UNIVERSITY TEACHERS: AN APPLICATION OF A THEORY OF EVALUATION AND AUTHORITY

Robert R. Hind, Sanford M. Dornbusch, and W. Richard Scott

A theory of evaluation and authority first put forth in 1967 seeks to explan the stability and effectiveness of authority systems in organizations.[1] Since then we have applied the theory to Roman Catholic priests, interns and residents in hospitals, teachers in public schools and alternative (free) schools, varsity football players, a physics research group, an electronics assembly line, the staff of a student newspaper, and workers in a factory and a hospital in Nigeria. These empirical studies tested portions of the theory and led to its reformulation.[2] This paper attempts to show how some aspects of the theory apply to the process by which a university faculty is evaluated.[3] Our empirical findings also have implications for directing change in a university.

Every organization is a power structure oriented toward controlling and coordinating activities. The structure may be hierarchical, with control exercised by bureaucratic superiors, or collegial, with control exercised by fellow professionals of approximately equal rank. Our studies found considerable variation in control structures, even among professional bureaucracies. For Roman Catholic priests, the hierarchy

---

Robert R. Hind is Director, Western Region, of the Academy for Educational Development, Palo Alto, Ca. Sanford M. Dornbusch is Professor of Sociology at Stanford and a Research and Development Associate of the Center. W. Richard Scott is Chairman of the Department of Sociology at Stanford and a Research and Development Associate of the Center.

[1] W. Richard Scott, Sanford M. Dornbusch, Bruce C. Busching, and James D. Laing, "Organizational Evaluation and Authority," Administrative Science Quarterly, Vol. 12 (June 1967), 93-117.

[2] Sanford M. Dornbusch and W. Richard Scott, Evaluation and Authority (New York: McGraw-Hill, forthcoming).

[3] Robert R. Hind, "Evaluation and Authority in a University Faculty" (unpublished Ph.D. dissertation, Stanford University, 1968).

assumes more importance as priests get older. Among public school teachers, hierarchical controls are weak and collegial controls are almost nonexistent. This paper presents evidence that at Stanford University--a professional bureaucracy--collegial control is much stronger than bureaucratic control.

Research Procedures

The Sample

In 1968 we interviewed a random sample (stratified by rank) of 100 faculty in the School of Humanities and Sciences at Stanford University. A school of this sort, often called the college of arts and sciences, is the core of most major American universities. The fields represented by these faculty members were similar to those in the traditional liberal arts college. Since our data were gathered, attempts have been made at Stanford to increase the emphasis on teaching in the evaluation process.[4] Thus, the results given in this paper may not depict the current situation at Stanford.

In any case, this paper is not meant to be descriptive. We hope our results are abstract enough to be applicable to a wide variety of academic institutions. Our propositions and hypotheses are intended to be generalizable. Therefore, it is appropriate to state here some of Stanford's characteristics so that the skeptical reader can have a basis for limiting the applicability of our findings.

Stanford University is an independent, nondenominational university in a northern California suburb. It has 12,000 students, about half undergraduate and half graduate students. Admission is highly competitive. About half of the 1,000 regular faculty are in the School of Humanities and Sciences from which we drew our sample; the rest are in six graduate and professional schools. These figures were very nearly the same in 1968.

---

[4] A follow-up study of their relative success is under way.

Stanford was recognized as a good regional university from its founding in 1885, until the late 1940's. An aggressive central administration then combined government research grants, gifts from private donors, and active recruitment of faculty to construct one of the nation's leading universities, with a strong research and graduate training program. As a result of this recruitment and expansion, most of the faculty in our sample had been at Stanford less than 10 years, although a sizable minority, 15 percent of our sample, had been at Stanford over 20 years. Younger men selected primarily for their research dominated the faculty. The organization of the faculty was, and still is, strongly departmental, with each department controlling its own instructional program. The central administrations of the University and of the School of Humanities and Sciences exercised their authority primarily through the allocation of resources to the departments.

An analysis of the educational process at Stanford helps to illustrate the decentralization of power at research-oriented universities, the permeability of universities--social systems whose sets of influential evaluators may be far beyond the individual campus--and the difficulties of redressing the imbalance between the emphasis on research and the emphasis on teaching. Clearly, Stanford is close to one extreme of university organization, in the United States, yet it is representative of some major institutions. It can be used as a baseline for comparison with studies of other universities, colleges, and junior colleges.[5]

The Interview

Each professor in the sample was asked a series of questions about the influence of the evaluation of four tasks in determining university rewards, and the identity and influence of various evaluators. Since Stanford has an annual review of all salaries, with no fixed scale for any rank faculty members readily recognized that evaluations directly and indirectly affected the reward system. By contrast, the public elementary and

---

[5]Charles E. Bidwell, "The School as a Formal Organization" in James G. March, ed., Handbook of Organizations (Chicago: Rand McNally, 1965); Burton R. Clark, The Open Door College (New York: McGraw-Hill, 1960); Talcott Parsons and Gerald M. Platt, "The American Academic Profession" (multilithed, 1968).

secondary school teachers whom we studied had difficulty conceiving of a system of rewards that had minimal variability and that depended on infrequent evaluations based on non-communicated criteria. Faculty members at Stanford knew that they were being evaluated and that evaluations mattered.

The questions were designed to elicit information about the individual experiences and perceptions of respondents, not what they thought about institutional processes in general. The list of questions was repeated for each of the principal tasks to which respondents attached inportance. Where strong disparities or dissatisfactions were noted, we probed for further details, noting these along with volunteered comments.

All interviews were conducted by the senior author of this paper, who was then Staff Director of the Study of Education at Stanford. The research reported here was part of a program of self-study. Only five faculty members, four full professors and one associate professor, chose not to participate in the study, and they were replaced by alternates also selected randomly. The noncooperating professors were not so numerous as to distort our results. The 34 full professors, 33 associate professors, and 33 assistant professors were representative of their ranks in the faculty of Humanities and Sciences.

Because the interviewer held an administrative position in the University, assurances of complete anonymity and confidentiality were conveyed in the letter inviting participation and in the oral introduction to the interview. We believe that these assurances were accepted, and our impression is supported by the large number of responses that criticized administrators and administrative procedure.

Applying the Theory

The University as a Professional Bureaucracy

In our theory, organizations are defined as power structures in which some participants are given differential influence over organizational rewards and penalties in order to control other participants. We believe that evaluations are necessary, but not sufficient, for the control

of behavior. We distinguish between the subjective importance of an evaluator or evaluation and the influence of the evaluator on organizational rewards and penalties.

If participants care about organizational rewards and penalties, then evaluators who have more influence over their distribution will be more important to participants than will evaluators who have less influence. Similarly, evaluations that have more influence over the distribution of organizational rewards and penalties will be more important to participants than will those that have less. Accordingly, to the extent that participants value organizational rewards, influential evaluators and evaluations will be perceived as important by the people being evaluated. Not all important evaluators need be influential, however. Our parents, our spouses, and our former teachers may be important to us even though they have no effect on what happens to us within the organization.

Therefore, given our previous research showing that influential evaluators become important and that influential evaluations are deemed important, we concentrated in this study on the extent to which evaluators and evaluations influence the distribution of organizational rewards and penalties. The bases for influence may be varied and the means of exerting influence may shift, but we are simply studying the end product: the impact on the distribution of organizational rewards and penalties. By this simplification considerable detail is lost, but comparability across studies and methodological simplicity are gained.

Data were collected on the evaluations of four faculty tasks: teaching, research (or scholarship), university service, and external service. The last two tasks were not seen by the professors as having much influence upon the distribution of university rewards and penalties. The evaluations of university service were perceived to be "very influential" or "extremely influential" by only eight percent of the sample, and external service by only three percent. When asked how much influence evaluations of these two tasks <u>should</u> have, the faculty showed no desire to raise their influence. We can, therefore, safely concentrate on the evaluation of only two tasks: teaching and research (with research including what some faculty call scholarship). These are the principal

tasks of the Stanford faculty. Research was considered "very influential" or "extremely influential" in the reward system by 78 percent of the faculty, teaching by 20 percent. The great emphasis on research, of course, provides a context for comparison with other institutions.

Professionals are oriented toward control by standards shared by members of their own occupational group. They attach greater importance to evaluations by their professional colleagues than to evaluations by non-professionals. Stanford University is a professional bureaucracy, and its structure of influence emphasizes collegial evaluations. We hypothesized that professional colleagues would be the most influential evaluators at Stanford--more influential than superiors in the administrative hierarchy.

Each professor was asked to name those persons who, for each task, made evaluations that might influence University rewards. Each professor was shown the following list from which to select influential evaluators.

1. Students
2. Department head
3. Department colleagues
4. Faculty members in other departments
5. Members of the same discipline in other institutions
6. Persons having a say in government or foundation grants
7. Other outsiders (please name)
8. Dean of the School and his staff
9. Appointment and Promotion Committee
10. Other faculty committees (please name)
11. Provost and his staff
12. President
13. Trustees

We then asked each professor how much each of his evaluators influenced the distribution of organizational rewards and penalties. Table 1 shows the perceived influence of evaluators for the two principal tasks, with some groupings of evaluators who are less frequently considered influential. Table 1 combines two kinds of influence, the influence of evaluations of each task and the comparative influence of groups of evaluators. We have already noted that evaluations of research are far more likely to be "very influential" or "extremely influential" than are evaluations of teaching. Therefore, no group of evaluators is likely to be high in influence when they are evaluating the task of teaching. If we combine the percentages labelled as "slightly infuential," "not at all influential,"

TABLE 1

Perceived Influence of Evaluators
(N = 100)

	Extremely Influential	Very Influential	Moderately Influential	Slightly Influential	Not At All Influential	Not Named
Teaching						
1. Students	6%	13%	17%	22%	5%	37%
2. Department head	17	18	12	5	4	44
3. Dept. colleagues	4	15	27	8	6	41
4. Other faculty	0	1	4	11	1	83
5-7. Outsiders	0	4	4	6	1	85
8. Dean and staff	3	7	2	9	2	78
9. A. & P. Committee	2	1	0	3	0	94
10. Other committees						100
11-13. Provost and above	1	2	1	1	0	95
Research or Scholarship						
1. Students	0%	0%	1%	2%	2%	94%
2. Department head	34	27	5	2	2	30
3. Dept. colleagues	33	28	12	5	3	18
4. Other faculty	0	2	9	6	2	81
5. Colleagues at other institutions	15	25	22	6	2	30
6. Grant makers	4	4	3	8	0	80
7. Other outsiders	3	2	1	2	1	91
8. Dean and staff	7	15	7	5	1	65
9. A. & P. Committee	7	4	1	4	1	83
10. Other committees						100
11. Provost and staff	1	5	6	1	1	86
12-13. President, Trustees	0	2	1	2	1	95

or "not named" at all, we have the extent to which each group of evaluators was considered low in influence. For teaching as a task, the majority of the faculty considered every evaluator low in influence. For research, on the other hand, there were three groups of evaluators who were not considered low in influence by the faculty. Therefore, when we consider the perceived influence of evaluators, we are looking only at the relative influence of evaluators of each task. Evaluators of research, in general, have more influence; yet we can learn much from comparing groups of evaluators with respect to the influence of their evaluations of teaching as well.

The data support the hypothesis that professional colleagues are more influential than administrative superiors. Table 1 shows that many more professors name their department colleagues as influential evaluators than name any administrator in the School or University. Administrators other than the department head are not only named less often, but they are also perceived as lower in influence than department colleagues. For example, only ten percent of the faculty saw the Dean and his staff as "very influential" or "extremely influential" evaluators for the task of teaching, and only 22 percent saw the Dean and his staff as "very influential" or "extremely influential" in the evaluation of research. The department heads, department colleagues, and students were all perceived as having more influence than the Dean for the task of teaching. For research, the department heads, department colleagues, and colleagues at other institutions were all rated higher in influence by the faculty than was the Dean. This lower influence of administrators is probably related to the combination of two perceptions: that administrators do not know what the appropriate evaluative criteria are, and that faculty performances are not visible to them.

Table 1 also shows that the department head, who is both colleague and administrative superior, has a critical role. His professional credentials serve to render his evaluations acceptable to his "colleagues," while his bureaucratic office requires him to gather certain kinds of information about the work of his colleagues, attempt to exercise certain kinds of controls over them, and help to decide how rewards and penalties

are to be allocated among them. It is the creation of special roles such as department head that leads us to conclude with Etzioni that "a large amount of control over professional performances has been transferred from the professional community to the professional organization."[6]

Since the faculty's emphasis on evaluations by professional colleagues is central to this paper, we examined the responses of faculty in the different ranks to see if there was any major variation in the perceptions of assistant professors, associate professors, and full professors. Assistant professors are rewarded by tenure, promotion, and salary increases; associate professors are given promotions and salary increases; and full professors receive only salary increases. These differences in available sanctions had no impact on our findings. The pattern of influence among evaluators was identical for the three ranks. For each rank, department heads, department colleagues, and students were all perceived as more influential than the Dean for the evaluation of both teaching and research. So, despite the differences in rewards and penalties, faculty members in each rank agreed on the relatively low influence of even the most influential administrator.

Visibility and Permeability

Our theory states that evaluations based on greater visibility of performances and outcomes are more likely to be considered soundly based by the participant. In addition, participants will want evaluations to be soundly based so that their performances can affect the evaluations.

We defined visibility as the frequency or proportion of performances observed or outcomes reviewed by an evaluator. This gave us four measures of visibility in our studies of nurses and public school teachers.[7] Each specific measure of visibility of work to an evaluator produced equivalent results. Among nurses and teachers, each measure of visibility was

---

[6] Amitai Etzioni, A Comparative Analysis of Complex Organizations, (New York: Free Press of Glencoe, 1961), p. 259.

[7] Gwen Marram, Sanford M. Dornbusch, and W. Richard Scott, "The Professionalism of Elementary School Teachers, Teaming, and the Visibility of Teaching" (tentative title; article submitted for publication).

strongly correlated with the perceived soundness of evaluations. The evaluators to whom work was more visible were considered to have more soundly based evaluations, and their evaluations were considered more important by the nurses and teachers. The only evaluators who were high in influence and lower in importance than would be expected were school superintendents and directors of nursing, to whom the work was not visible.

We did not directly collect data on visibility from the Stanford faculty, but the impact of visibility is clear. Since Stanford is a professional bureaucracy with the current evaluative structure close to the preferences of the faculty, visibility affects the current system of evaluation. Students, for example, were seen as major evaluators of teaching. Students were the most often named evaluators for teaching, just behind the department head and tied with department colleagues in the reported influence of their evaluations. However, even this relatively large perception of student influence in the evaluation of teaching actually underestimates student participation in the evaluation process. In a university, teaching is often visible only to students. Therefore, students are the primary source upon which other evaluators base their judgments of teaching. When they were asked the sources of information used by department heads and department colleagues in the evaluation of teaching, the faculty reported that students were the source of information for 81 percent of the department heads and 93 percent of the department colleagues.

Yet the high visibility of teaching performances to students conflicts with the faculty's emphasis on evaluations by professional colleagues. Although visibility increases the likelihood of sound evaluations by students, their ignorance of professional standards reduces faith in their judgment. Faculty members were divided on the subject of students as good evaluators of teaching. We asked no specific questions about the quality of any of the evaluators, yet over half of our sample commented on it with regard to students. Twenty-eight percent volunteered comments which indicated that students were good evaluators of teaching, while almost an identical number, 29 percent, said that students were poor or

unreliable evaluators of teaching. This finding is in sharp contrast to that of a related study of nurses and teachers. Among nurses and teachers, there is a very high correlation between the visibility of their performance to an evaluator and the desired influence of that evaluator. But the evaluators in these studies were professionally trained nurses, head nurses, teachers, and principals.

We defined organizational evaluators as those persons who influence the distribution of organizational rewards and penalties. Organizational evaluators need not be within the traditional boundaries of the organization. The greater the visibility of task performances or outcomes to outsiders, the more likely it is that the boundaries of the organization will be permeable and that outside evaluators will influence the distribution of organizational rewards and penalties. Research is visible to colleagues at other institutions through the publication of results, whereas teaching is not. Seventy percent of the faculty named colleagues at other institutions as organizational evaluators of research and 40 percent saw these distant colleagues as "very influential" or "extremely influential." But since teaching is not highly visible to outsiders, the organization is less permeable with respect to the evaluation of teaching. Only 15 percent of the faculty named any outsiders as organizational evaluators of their teaching, and just four percent saw them as "very influential" or "extremely influential." If the two sets of evaluations are combined, we find that 19 percent of all evaluators named (for both research and teaching) were located outside the organization whose rewards and penalties were being distributed. Yet these 19 percent were actually influencing the allocation of those sanctions. The permeability of the University reported for research is striking because we are considering only influence on the distribution of Stanford's rewards, not outside honors and awards. Outsiders' influence may be indirect, via competing offers or reflected reputation within Stanford, but it nevertheless influences Stanford's reward system.

Younger faculty were particularly likely to emphasize those tasks--research and scholarship--that would be most visible to their colleagues at other institutions. At Stanford, as at many colleges and universities,

there had been much public discussion about the relative inattention to teaching among the faculty prior to our study. It was expected that this shift in values would be reflected in more concern about teaching among younger professors. On the contrary, younger faculty tended to be more mobile and to care more about participating in a national and international job market. The visibility of research to other institutions was so powerful a force that the increased emphasis on teaching was nowhere evident for younger faculty, if it existed at all. We examined our data to see if this result was spurious. However, within each rank, younger faculty emphasized research more than older faculty. Younger professors appear to be emphasizing the task for which the organizational boundaries are most permeable.

The permeability of the university to outside evaluators is also illustrated by the influence of the awarders of grants in the evaluation of research. The influence of grant makers is not very high, with only eight percent of the faculty viewing them as "extremely influential" or "very influential" evaluators of research. Nevertheless, the perceived influence of grant makers did vary with the amount of money received by each department. For departments which received less than $1,000 per year per professor in grants, 92 percent assigned no influence to grant makers in the evaluation of research, and none saw them as "very" or "extremely" influential. For departments receiving from $1,000 to $9,999 per year per professor, 80 percent saw grant makers as without influence and only eight percent saw them as high in influence. For departments that received $10,000 or more per year per professor in research grants, the proportion seeing grant makers as not influential declined to 69 percent, and the proportion seeing them as "very" or "extremely" influential rose to 17 percent. Thus, although the proportion who perceived grant makers as high in influence was never large, the amount of funds received by a department was directly related to the perceived influence of grant makers in that department.

Dependence Upon a Body of Theory

It was hypothesized that satisfaction with the evaluation process would be related to perceived agreement among evaluators; that is, if evaluators are seen as agreeing, the professor is less likely to be dissatisfied with their evaluations. To measure the extent of perceived agreement we asked, "How much agreement do you think there is among people in your discipline in their evaluations of (task) done by their colleagues?" Somewhat surprisingly, agreement among evaluators as perceived by the faculty was approximately equal for the tasks of research and teaching. There were only minor variations among fields, except for the arts, which tended to be low in perceived agreement among evaluators for both teaching and research. There was also a slight tendency for faculty in the natural sciences and mathematics to perceive a higher level of agreement among evaluators for research than for teaching.

An analysis of the relationship between perceived agreement among colleagues' evaluations of research and satisfaction with evaluations by those colleagues reveals a gamma of .59 between high agreement and high satisfaction. Eighty-three percent of those who perceived high evaluator agreement among colleagues were high in satisfaction with evaluations, compared with 56 percent of those who perceived low agreement among their colleagues.

Our theory views satisfaction as partially the product of perceived agreement among evaluators, but clearly an interpretation emphasizing cognitive consistency would also fit the data. Such a cognitive interpretation would portray satisfied faculty as more likely to perceive agreement. Our theory specifies a direction from agreement among evaluators to satisfaction among faculty, and the next set of findings indicates some basis for this view.

Greenwood, among others, has noted that one of the criteria for a profession is its dependence on a body of theory.[8] We therefore hypothesized that, in a professional group, agreement among evaluators would be

[8]Ernest Greenwood, "Attributes of a Profession," Social Work, Vol. 2 (July 1957), 44-55.

more likely when the discipline was perceived to be based upon a body of theory. The body of theory would provide a common set of standards for evaluation.

To get a measure of dependence upon a body of theory, we asked, "Some academic fields or disciplines are seen as strongly based on a systematic body of theory which serves as the basis for research hypotheses and dominates teaching and scholarship. Other fields are less reliant on a central body of theory. To what extent does your own field have a central body of theory to guide research, scholarship, and teaching?" For the 77 respondents who answered both questions, we obtained a gamma of .45 relating high dependence upon theory and high perceived agreement among the evaluators of their research. Sixty-four percent of faculty who saw their discipline as "extremely dependent" or "very dependent" upon theory perceived high agreement among their evaluators. Only 40 percent of those who felt their discipline had a low dependence upon theory perceived high agreement among their evaluators. Dependence upon theory is more prevalent in the natural and social sciences than in the humanities. None of the natural scientists, 17 percent of the mathematical scientists, and 29 percent of the social scientists, saw their field as "slightly dependent" or "not at all dependent" upon theory, compared to 51 percent of the faculty in the humanities.

From this finding, we predicted that overall satisfaction with the evaluation process would be positively related to perceived dependence upon theory within a discipline. Since many respondents gave middle responses, we trichotomized the responses according to dependence upon theory. We found that overall satisfaction with evaluations for those professors in disciplines perceived to be highly dependent upon theory was 60 percent; for those in disciplines with medium dependence upon theory it was 52 percent; and for those in disciplines with low dependence upon theory it was 38 percent. These data support the view that satisfaction with the evaluation process is related to perceived agreement among evaluators, which is in turn related to perceived dependence upon a body of theory. This finding appears to be the basis for the professional emphasis upon collegial evaluations; that is, evaluations based upon a shared body of knowledge.

Our studies of nurses and public school teachers provide supporting evidence for the relationship between a shared body of knowledge and collegial control. Teachers, who overwhelmingly consider "personality" more important than training in producing a good teacher, are reluctant to permit evaluations by colleagues to be influential. Nurses, who consider their training vital, are oriented to the evaluations of other nurses.[9] Thus, it appears that a shared body of knowledge is a major basis for the collegial control that is presumed central for professionals.

Effort and the Evaluation Process

It has already been argued that if participants care about organizational rewards and penalties, then evaluations that have more influence over the distribution of organizational rewards and penalties will be more important to participants than will those evaluations that have less influence. The data from our study of university faculty partially support, and partially do not support, this proposition. The effect of the organization's reward structure upon faculty effort is partly vitiated by the internalization of professional standards, which remain powerful and independent of the evaluation-reward system. The Stanford faculty recognized that evaluations of research were far more influential than evaluations of teaching. Seventy-eight percent saw evaluations of research as "very influential" or "extremely influential," compared to only 20 percent for teaching. Yet most of the faculty reported that they spent more time on the various forms of teaching than on research. This finding appears to defy the attempts of the evaluation system to control the allocation of effort. Thus, the discrepancy does produce strain and a desire for better balance.

First, 40 percent of the faculty wanted to increase the amount of time devoted to research, compared to only 16 percent who wanted to reduce their research effort. A substantial majority of the faculty wanted to continue with the amount of time currently devoted to teaching, with slightly more

---

[9] Marram, Dornbusch, and Scott, op. cit.

wishing to reduce their teaching time than wishing to increase it. These desired shifts in allocation of time would slightly increase time for research and slightly reduce time for teaching, a preference that would, on the average, produce a somewhat better balance between effort and reward.

Second, the faculty wanted to increase the influence of evaluations of teaching. Fifty-one percent of the professors felt teaching should be "very influential" or "extremely influential," a marked increase. Yet an even higher proportion, 67 percent, thought research should be "very" or "extremely" influential. Thus, although the percentage of faculty who felt that teaching should be highly influential is more than double the number who now see it as high in influence, the decline in the desired influence of research is not so great as to make teaching and research equal in influence. The faculty viewed research as the most influential task and believed that it should continue to be so. The desire for balance is indicated, however, by the strong desire to increase the influence of evaluations of teaching. Fifty-three percent of the faculty wanted to increase the influence of teaching upon rewards, and only two percent wanted to decrease it. Faculty members were seeking to balance their expenditures of effort with the influence of their tasks upon rewards.

Teaching seems to have been internalized as part of a professional conception of an appropriate professional role. Despite this feeling of professional responsibility for teaching, there was, as predicted, a tendency among faculty members to spend more time on those activities which were believed to benefit them most. Among those who saw teaching as "moderately," "very," or "extremely" influential, 65 percent were above the median on time spent teaching undergraduates and 77 percent were above the median on time spent on graduate teaching. For those who saw teaching as "slightly influential" or "not at all influential," only 43 percent were above the median on time spent on undergraduate teaching and 47 percent above the median on time spent on graduate teaching.

The same pattern was found for research. Those who saw research as highly influential in the evaluation and reward system were more likely to be above the median in time spent on research, compared to those who saw

research as less influential; the respective proportions are 63 percent and 40 percent. Once again, there appears to be an interaction between professional responsibility, internalized standards for effort, and a tendency for faculty to be guided by their perception of relative reward.

Implications for Changing the Evaluation System

We have argued that the University is a professional bureaucracy whose evaluation process is dominated by professional colleagues. How, then, could the University have an evaluation system that differed markedly from the one desired by its faculty? One might assume that the faculty and the central administration were in opposition with respect to producing a more balanced evaluation process, but that simply was not true. The public statements and private efforts of the administrators of Stanford were, and are, directed toward increasing the influence of evaluations of teaching. Thus, it appears that everyone wanted change, but waited for some other agent to bring it about.

The visibility of research in a national and international market and the lack of visibility of teaching to persons outside the universiry were probably major factors in preserving the existing reward system. Since Stanford faculty will continue to participate in a broader social system than the University, and since the University's boundaries for the evaluation of research are permeable, the faculty and the administration must take these realities into account in any future attempts at change. But faculty dissatisfaction with the relative de-emphasis of teaching in the evaluation system provided an opportunity to redress the imbalance between effort and reward.

Even if the administration of the University presses for change, the change itself must be carried out by the faculty--the faculty has the power to bring about a change or thwart it. The evaluation-reward process appears to be the primary means by which leadership groups in the administration and the faculty can initiate change when the change is in a direction approved by a majority of the faculty. If faculty effort on any task is desirable, that task should receive appropriate emphasis in the evaluation-reward system.

Under the auspices of the Dean of Undergraduate Studies, a new official charged with improving some aspects of faculty teaching, we will survey the Stanford faculty to examine the effect of recent attempts to change the evaluation system by increasing the weight given to teaching. The study reported here provides the baseline for the measurement of change.